

# Weekly Report (2018.1.1-2018.1.7)

TASK	DEADLINE	CURRENT PROGRESS
可视交互引擎论文	2017.12月底	目前进展：补充了沉浸式可视分析架构与平台 存在问题：时间不够充裕 计划：提高效率，与其他同学分工合作
AI课程论文	2018.1.10	目前进展：完成课堂报告，综述报告确定框架与基本内容 计划：阅读相关论文，撰写综述报告
游戏设计与开发课程项目	2018.1.15	目前进展：基本完成 存在问题：WebGL相关细节调整 计划：熟悉 WebGL 用法，修饰细节
VR课程项目	2018.1.23	目前进展：安装学习了Unity with IOS 计划：学习 Unity with IOS 的3D图形渲染，实际进行开发实验

## Done

### 1. Paper writing.

Add content to the subsection that Methods and Techniques of intelligent visual analysis. Due to the lack of time, the addition is not too much. Next week I will write with Kong to raise efficiency. Just sit down and spend more time to do research. Everyone is equal in the face of time.

### 2. Paper reading.

How Do Ancestral Traits Shape Family Trees over Generations?

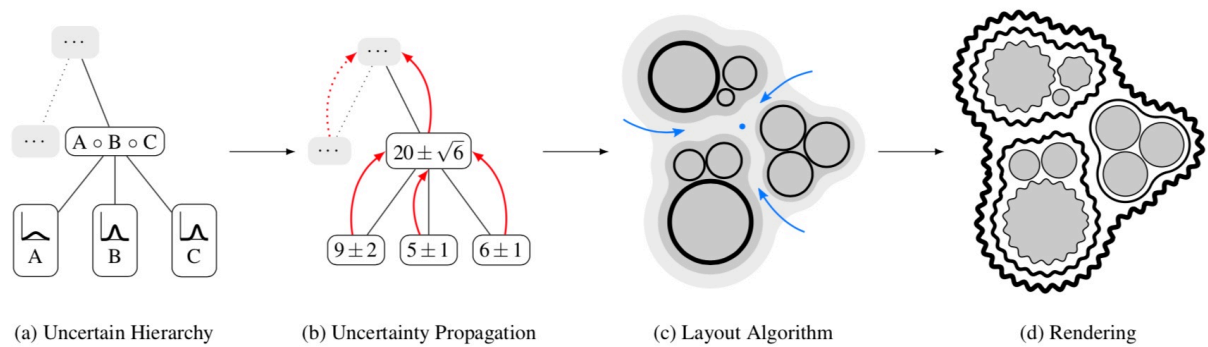
This work is based on multi-generational datasets. They develop a system TreeEvo that assists users to generate and test empirical hypotheses for multi-generational research. They make use of Sankey diagram and treemap, and the interaction design seem great.

### 3. Group meeting report.

Title: Bubble Treemaps for Uncertainty Visualization.

The authors are Germen. Their contribution is threefold: First, they propose a layout based on circle packing to use space purposely, achieving a reasonable trade-off between a compact representation of the hierarchy and its inherent information. Second, they define node contours analytically, resulting in a new parameter domain to be used for additional visual variables, in particular, for uncertainty visualization. Third, they describe different models of uncertainty and discuss their relation to hierarchical data.

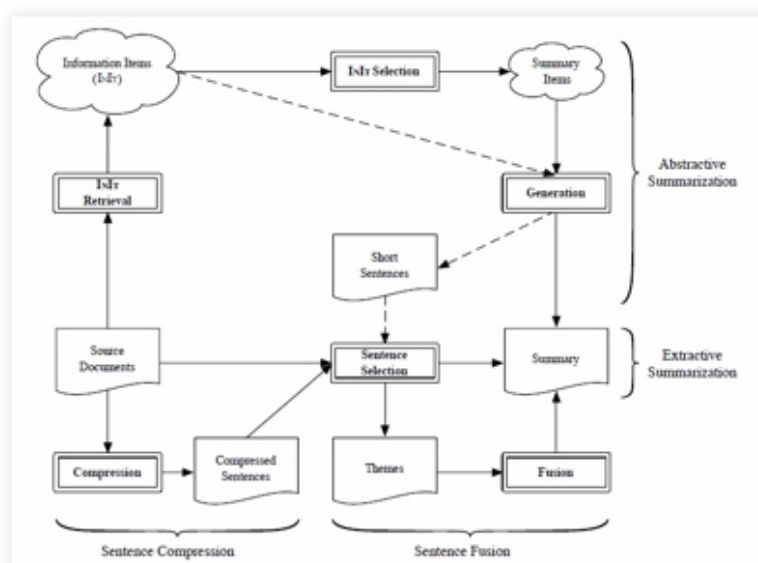
The overview is as follow:



4. AI course presentation.

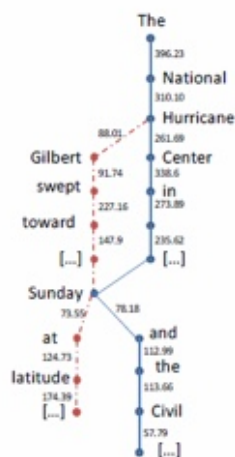
## 生成式文本摘要方法

- 通过文本到文本生成文本摘要
  - 提出了一种基于信息项（INIT）概念的方法，信息项是文本或句子中相关信息的最小元素。
  - 四个操作步骤：INIT检索，句子生成，句子选择和摘要生成。
  - 试图控制文档的内容和结构。
  - 评测结果：TAC 2010数据集的评估结果令人满意，金字塔分数为0.315，语言质量为2.174，总体响应度为2.304。



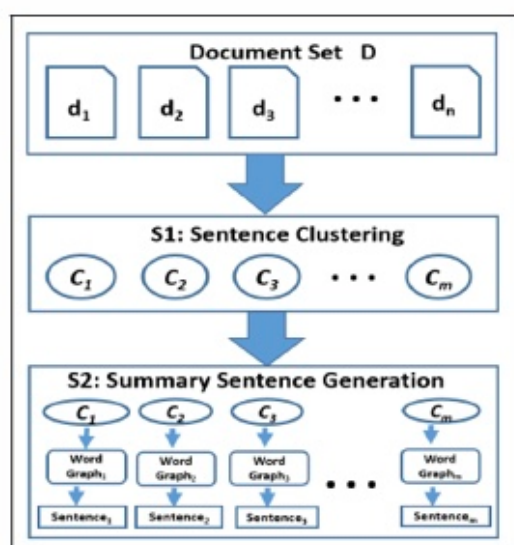
## 生成式文本摘要方法

- 通过使用词图生成文本摘要
  - 操作步骤：压缩和合并来自句子的信息以形成新的句子 → 采用抽取式文本摘要方法COMPENDIUM来确定应该选择哪个新句子来生成文摘。
  - 实验证明，通过结合抽取式和生成式，可以获得更高质量的摘要。
  - 评测结果：在DUC 2002获得0.405的ROUGE得分



## 生成式文本摘要方法

- 通过基于ILP的多句子压缩生成多文档的文摘
  - 操作步骤：从多文档集合中选择最重要的文档
    - 来自最重要的文档的每个句子被用来生成单独的集群
    - 与集群句子具有最高相似度的其他文档的句子被分配给该集群
    - 通过由每个聚类的句子形成的一个单词图结构，生成K个最短路径
    - 采用ILP从最短路径集合中选择句子
  - 优点：最大限度地提高信息内容和语言质量，减少最终总结中的冗余。
  - 评测结果：DUC2004和DUC2005数据集，本方法的ROUGE得分比两个数据集上的最佳抽取摘要都好，并且该系统胜过基于多句子压缩的生成式文摘。

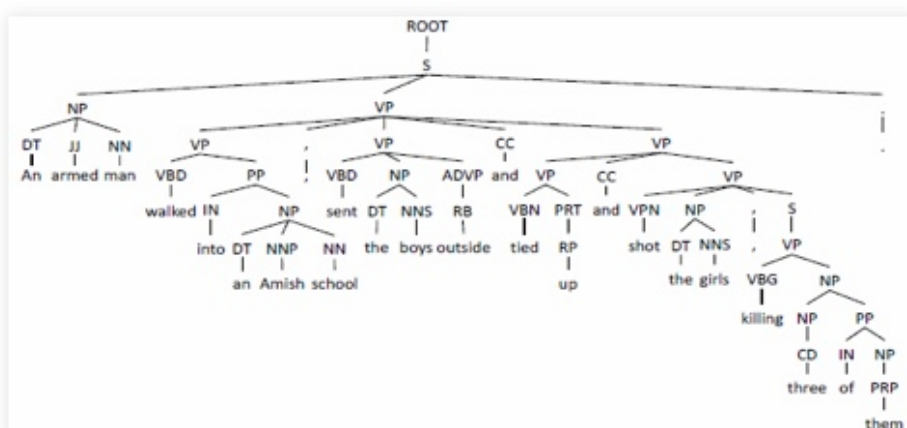


## 生成式文本摘要方法

- 基于短语选择和合并生成多文档的文摘
  - 通过探索比句子更细粒度的名词或动词短语的句法单位来创建新的句子
  - 操作步骤：从输入文档中提取由名词或动词短语表示的概念和事实 → 利用文档内容的冗余为每个词组计算显著性分数 → 选择并合并短语，以得到有效的新句子
  - 评测结果：使用自动化的金字塔评估指标对TAC 2011数据库进行实验评估，在0.6和0.65的阈值分别为0.905和0.793，比TAC 2011中的其他系统要好。



Visual Analytics Group | CAD&CG 国家重点实验室



- ROUGE 系统
  - 基本思想：将待审摘要和参考摘要的n元组共现统计量作为评价依据，然后通过一系列标准进行打分。
  - 标准包括： ROUGH-N、ROUGH-L、ROUGH-W、ROUGH-S和ROUGH-SU几个类型。

5. AI course survey report.  
Determining the writing framework and basic content.

### TODO

1. Write paper.
2. Finish the AI course survey report.
3. Set out to the Game design and development course project.